Placebo Insight: The Rationality of Insight-Oriented Psychotherapy

David A. Jopling
York University

It is widely believed that the insight-oriented psychotherapies provide their clients with valid methods of self-exploration that lead to bona fide self-knowledge. It also is widely believed that clients' insights must be true in order to be therapeutically effective. Both these claims are implausible. I argue that because clients face significant epistemic pressures in the therapeutic encounter, the insight-oriented psychotherapies are highly susceptible to generating placebo insights, that is, illusions, deceptions, and adaptive self-misunderstandings that convincingly mimic veridical insight but have no genuine explanatory power. The insight-oriented psychotherapies also are highly susceptible to generating therapeutic artefacts that appear to confirm the insights acquired by clients. The powerful treatment methods to which clients are subjected generate some of the very psychological and behavioral facts that clients claim to “discover” in their explorations. This impugns the scientific status of the insight-oriented psychotherapies. © 2001 John Wiley & Sons, Inc. J Clin Psychol 57: 19–36, 2001.

Keywords: insight; self-knowledge; the concept of truth in insight-oriented psychotherapy; insight-mimicking explanatory fictions; therapeutic conformability; therapeutic artefacts; the concept of placebo

It seems obvious that the insight-oriented psychotherapies are distinguished from the behavioral and brief psychotherapies by virtue of the emphasis they place on the exploration of the soul and its depths. It also seems obvious that the clients of insight-oriented psychotherapies who engage in this lengthy and arduous exploratory work should, and in

Correspondence concerning this article should be addressed to: David A. Jopling, Department of Philosophy, York University, 4700 Keele St., Toronto, Ontario, M3J 1P3 Canada; e-mail: jopling@yorku.ca.

The term insight-oriented psychotherapy is used here as a generic term to cover the so-called talking curees: for example, the various forms of Freudian and neo-Freudian psychoanalysis, Jungian analysis and its offshoots, Gestalt psychotherapy, person-centered psychotherapy, and the various forms of existential psychotherapy, among others.
fact do, emerge edified, and with a greater clarity about themselves than they had at their journey’s outset. Finally, it seems obvious that their newly won clarity somehow improves their lives, or at least serves the ends of psychological healing.

These appear to be good assumptions, not only because the clinical evidence supplied by insight-oriented psychotherapists and the self-reports of satisfied clients seem to bear them out, but because it is hard for us to imagine what else the insight-oriented psychotherapies could possibly be if they did not involve the exploration of the soul and its depths. The link between self-exploration, insight, and healing seems natural and obvious to us. The link is further reinforced by the persuasive metaphors called upon by insight-oriented psychotherapists to characterize their practice: For example, the insight-oriented psychotherapies are said to hold up a mirror to the soul, to proceed by a kind of archaeological excavation of hidden layers of psychological sediment, and to result in clients “getting in touch” with an “inner” or “core” or “authentic” self.

But is this ostensibly obvious link credible? Why assume in the first place that what occurs in the insight-oriented psychotherapies is exploration and discovery rather than something that looks like exploration and discovery? Why assume that the exploration of the soul—rather than one’s society, or one’s body, or one’s past lives—is essential to healing? And why assume that the kind of exploration that occurs in the therapeutically successful cases of insight-oriented psychotherapy is accurate and truth-tracking, and that it actually succeeds in uncovering the deepest layers of the soul (or, depending on the therapeutic modality, the deepest layers of the unconscious, the core self, or the existentially authentic self)?

These are philosophical—and specifically epistemological—questions. They are to be distinguished from the kinds of questions that could be answered by appeal to empirical data (e.g., by accumulating more case studies and by pursuing more methodological fine-tuning): for it is not as if once all the data are in, the epistemological questions will take care of themselves. To cast light on these questions, and to coax them into a more manageable format, I focus on one specific question which has an obvious bearing on the scientific status of the insight-oriented psychotherapies: Does truth matter in the insight-oriented psychotherapies—and if so, to what extent? In other words, is it the case in the insight-oriented psychotherapies that clients’ explorations, discoveries, and insights, as well as psychotherapists’ interpretations of their clients, must be true in order to be therapeutically effective? To what degree are pseudo-insights, confabulations, and falsehoods—in short, explanatory fictions—tolerated as therapeutically beneficial?

Despite disagreement about the nature and accessibility of truth, most of the insight-oriented psychotherapies defend some form of commitment to the ideal of truth; more primitively, most of them defend some form of distinction between truth and falsity. It is these two commitments that (in their own view) help to distinguish what they do from charlatanism and pseudo-science. If they were to give up these commitments, and thereby allow that their treatment methods and the theoretical frameworks in which they are imbedded were not instrumental in getting at the truth, and not thereby instrumental in occasioning therapeutic change, then there would be no way of telling them apart from bogus methods and bogus theoretical frameworks that make similar-sounding claims. It is commitment to truth that is one of the crucial planks in the insight-oriented psychotherapies’ project to be scientific.

But these two commitments are jeopardized by a number of unwarranted inferential leaps in both the first-order practices and the second-order theories of the insight-oriented psychotherapies. To identify these, I begin with a moderately strong generic model of insight-oriented psychotherapy, using it as both guide and foil in order to address the question of whether truth matters. This strategy, focusing on the general form of insight-
oriented psychotherapy rather than on particular instances of it, will of necessity overlook the theoretical and methodological idiosyncrasies that serve to differentiate the many versions of insight-oriented psychotherapy from one another.

The Standard View

Despite enormous theoretical and methodological differences (Wallerstein, 1995), most of the insight-oriented psychotherapies agree on certain core principles: exploratory validity, therapeutic specificity, interpretive agency, and therapeutically effective insight. Together, these constitute what I call the standard view of insight-oriented psychotherapy.

According to the principle of exploratory validity, insight-oriented psychotherapy provides clients with a valid method of self-exploration that leads ultimately to bona fide self-knowledge. Whatever the particular theoretical orientation—psychoanalytic, Jungian, Gestalt, existential, feminist, or humanistic—insight-oriented psychotherapy is believed to afford clients the opportunity for discovery of facts about themselves that exist antecedently to therapeutic intervention, and that are logically independent of the theoretical and linguistic apparatus by means of which they are discriminated, identified, and described. These facts will speak, as it were, as long as the right treatment method and theoretical framework are brought to bear upon them.

Of course, precisely what counts as a relevant fact in the first place is a highly disputed issue that varies from one insight-oriented psychotherapy to another. In classical Freudian psychoanalysis, for example, the facts that are the object of therapeutic exploration range across such things as neuroses, complexes, and unconscious drives. In existential psychotherapy, which rejects psychoanalytic determinism and the theory of the unconscious, the facts that are the object of therapeutic exploration range across such topics as the client’s way of being in the world, and the client’s choices, projects, and stance toward death, and so on. Despite these differences, however, most of the insight-oriented psychotherapies would agree that there is something factual rather than fictional or artefactual—that is, something given, determinate, and prior to psychotherapeutic intervention—that it is the object of therapy to explore.

According to the principle of therapeutic specificity, the precise application of a therapeutic method has exact and nonsuggestive effects on the target disorder. The treatment method deals with specific problems that cannot be accessed by any more suitable means; and it engages special mechanisms whose operation vis-à-vis the target disorders are like those of a key to a lock, with the uniquely molded contours of the method fitting precisely onto the contours of the target disorders and exerting upon them an exact causal grip. This can be formalized as follows: Therapy T is differentiated from therapies XYZ because the precise and controlled application of T’s characteristic factors generates specific, nonplacebo, nonsuggestive effects in remedying the target disorders that are identified under T. For example, if a Client A is diagnosed as suffering from Disorder P (e.g., a neurotic obsession) according to theoretical Framework Y (e.g., classical Freudian psychoanalysis) and is treated with Method N (e.g., analysis and free association) for Time t1-t2, then Disorder P will be reduced because of the precise application of the characteristic factors of the Treatment N.

According to the principle of interpretive agency, which is a variant of the principle of therapeutic specificity, one of the high points of the therapeutic process (from the therapist’s point of view) is the development of an interpretation which uncovers, or at least points to, the truth about the client’s psyche and target disorders. Interpretations are regarded as powerful instruments of therapeutic change. They serve to guide, challenge, or reframe clients’ explorations in a manner that is ostensibly nonsuggestive and non-
manipulative; they render otherwise puzzling experiences intelligible and coherent; and, most importantly, they assist clients in acquiring veridical insight. The development of an interpretation is exceptionally difficult work, occurring at a critical stage in the therapeutic process. It requires searching for hidden clues and patterns among the massive amount of clinical material generated by the therapeutic intervention, and it requires judicious restraint from overinterpretation (i.e., making the clinical material more streamlined, coherent, or continuous than it in fact is).

What makes one interpretation true and another false? In practice, a number of singly necessary and jointly sufficient truth conditions are called upon. An interpretation is considered to be true if (a) it tallies with the facts of the client’s psyche, target disorders, and life processes; (b) it is followed by significant changes in the client’s conduct; (c) the client accepts it as true; (d) it moves the process of exploration forward by triggering new discoveries and opening up new topics; (e) it is analogous to and consistent with a sufficiently robust number of other case histories; and (f) it makes previously unintelligible experiences in the client’s life intelligible. There are, obviously, problems with each of these criteria and with their compatibility with one another, but these will not be addressed here.

The fourth core principle that is defended by most of the insight-oriented psychotherapies is the principle of therapeutically effective insight. From the client’s point of view, the high point of the psychotherapy is the acquisition of insight, the moment (or stage) when the client finally achieves an action-guiding understanding about what is going on behind the choices, behaviors, and feelings that have hitherto been so troubling and hard to understand (i.e., the behaviors identified under the target disorder). Most of the insight-oriented psychotherapies assume that insights are therapeutically effective if and only if they tally with the historical and psychological facts, and that inexact or incomplete insights that fail to tally with the facts are ineffective, or only temporarily effective (Kris, 1956). Truth-tracking insight, in other words, is a necessary condition for therapeutic improvement. Those clients who have acquired insight are considered to be well on the way to achieving a significantly higher degree of self-acceptance, emotional maturity, and self-responsibility than before they began therapy.

Naturally, the achievement of insight is an intense experience (or series of experiences) that comes after months of struggle and hard work. It is all the more intense because it occasions a certain disenchantment with pretherapeutic self-understandings, which, from the therapeutic and post-therapeutic point of view, appear to be superficial or misleading. The intensity of the experience is often enough to convince clients that the truth has finally been achieved—but this is clearly an unworkable criterion of truth: Intense feelings also can be generated by illusions and importantly false beliefs that mimic genuine insight. The criteria called upon to establish the truth-value of insights are similar to those used to establish the truth value of interpretations. Insights are considered to be true if (a) they tally with the facts of the client’s psychology and life processes, (b) they are followed by significant changes in the client’s conduct, (c) the client is inwardly affected by them, (d) they move the process of exploration forward, (e) they are analogous to and consistent with insights in other case histories, and (f) they make previously unintelligible experiences intelligible.

The standard view of insight-oriented psychotherapy can be summarized as follows: (a) Insight-oriented psychotherapy is a valid method of personal discovery that allows clients to discover truths about themselves, and to acquire bona fide self-knowledge; (b) the methods of the insight-oriented psychotherapies have specific and nonsuggestive effects; (c) one of the primary agents of therapeutic change in the insight-oriented psychotherapies is the therapist’s use of interpretations; and (d) the client’s acquisition of truth-tracking insight is a necessary condition for therapeutic improvement.
Problems with the Standard View

Stronger and weaker versions of the standard view are held by a great many practicing psychotherapists, but they are often held uncritically, without attention to the logical and epistemological issues they naturally raise. These issues would not be so troublesome if the standard view were intended only as a normative description of an ideal psychotherapeutic situation rather than as an account of the central structural features of actual clinical practice. As a normative ideal, it would set a standard for the practices of psychotherapy to emulate rather than describe the general structure of existing practices. Any logical or epistemological problems encountered at the clinical level could then be regarded as flaws in the interpretation or implementation of the ideal rather than as flaws in the ideal itself. However, it is doubtful that such a weak version would be acceptable to the majority of practicing insight-oriented psychotherapists.

The standard view contains a number of problematic epistemological assumptions. These problems are not to be confused with the more well-known empirical criticisms of the effectiveness of the insight-oriented psychotherapies, namely, that the insight-oriented psychotherapies are either ineffective in bringing about therapeutic change, or no more effective than credible placebos; and a fortiori, that insights are epiphenomenal and fail to make real differences to behavior (Eysenck, 1960; Prioleau, Murdock, & Brody, 1983). These criticisms may be true (or false, or overstated, or conceptually muddled), but they are independent of the epistemological criticisms.

First, the mere acquisition of insight is not a guarantee that the insight is true. This is the case even if the acquisition of insight is the culmination of months of hard work and struggle in a setting that constantly reinforces the image of insight-oriented psychotherapy as a legitimate venue for exploration and authentic discovery. A number of plausible alternative explanations of the acquisition and content of insight would have to be ruled out before the principle of therapeutically effective insight could be taken as a valid explanation of the phenomena. For example, what might appear to the client to be veridical insight may in fact be the product of a pervasive and undetected strategy of self-deception that finds support within the therapeutic setting, or it may be a false realization which becomes apparent with further (post-therapeutic) self-inquiry and further (post-therapeutic) consideration of the evidence.

Second, the client's level of conviction about the validity and authenticity of a newly won insight is not a guarantee that the insight is true. The client may be deeply convinced about—even form strong identifications with—an insight that is in fact psychologically and historically false. The client's level of conviction simply may be a function of a temporary lapse of judgment that is the result of systematically distorted epistemic and interpretive standards brought about by prolonged exposure to the epistemic and interpretive standards of the therapy. Naturally, this is not to deny the methodological significance of emotional arousal as a key ingredient in the acquisition of insight; but the fact that clients become emotionally aroused in certain stages in the therapy, during which they appear to arrive at an understanding of issues in their lives that they would not otherwise have understood, is not a guarantee that what they have understood in the emotionally aroused state is true.

Third, the therapist's conviction about the authenticity of the client's explorations, and the truth-value of the client's insight, is not a guarantee of the truth of the insight. Therapists have neither privileged nor authoritative access to their clients' life histories and psychologies, even when their therapeutic interventions have all the marks of being carried out correctly according to the standards of the therapeutic theory and the clinical norms of the community of practitioners. Nor do therapists occupy the point of view of
unbiased and impartial observers: They have vested interests in seeing their work succeed, they operate with theoretical orientations that have the potential to blind them to salient psychological and behavioral evidence, and they have only a finite amount of clinical material with which to work. Furthermore, neither consistency with previous case histories nor consensus among similarly trained therapists about the authenticity and validity of clients’ insights are guarantees that clients have come upon the truth.

Fourth, the occurrence of therapeutic change following the acquisition of insight is not a guarantee of the insight’s truth. To assume otherwise is to commit the fallacy of post hoc ergo propter hoc. There are a number of plausible alternative explanations that would have to be ruled out before accepting that therapeutic change was directly or indirectly a function of the truth of the insight. The change, for example, may have occurred because of factors less related to the truth-value of the insight than to its capacity to persuade the client with its apparent coherence, elegance, and explanatory power. One of the factors common to all forms of psychotherapy is that clients are supplied with a coherent rationale that explains their problems, gives otherwise puzzling symptoms a name (Torrey, 1986), and provides a socially sanctioned method of treatment. But even if the rationale is false, and the symptomatology and nosology unsubstantiated, the therapy may still be effective. In these cases, therapeutic agency rests on the client’s belief in the rationale’s validity and veridicality—actual truth value notwithstanding (Frank, 1989; Frank & Frank, 1991).

None of these four criticisms take issue with the clinical material which the standard view attempts to explain: that clients in the insight-oriented psychotherapies appear to make important discoveries, that the experience of discovery is intense, and that it sometimes is followed by therapeutic change. The clinical material is not in question; what is, however, is the interpretation it receives and the putative empirical support it provides for the validity and scientific status of the insight-oriented psychotherapies. The overlapping necessary conditions claimed by the standard view—that a properly conducted psychotherapy is a necessary condition for the acquisition of truth-tracking insight, which is a necessary condition for therapeutic change—are called into question by the fact that therapeutic explorations are not always authentic, that insights are not always veridical, and that therapeutic changes are not always the result of the precise application of the treatment methods.

This is not to say that it is logically impossible that the standard view holds in some cases—it simply does not hold with the degree of strength and the universality with which it is claimed to hold. Thus, before accepting the standard view as a valid account of the essential or necessary features of the insight-oriented psychotherapies, a number of equally plausible alternative explanations of the phenomena covered by the standard view first need to be ruled out. Clearly, the standard view’s account of the clinical material would be greatly enhanced if it were supported by controlled clinical and experimental studies rather than the uncontrolled clinical case studies that form the bulk of the relevant clinical evidence. While this would not conclusively establish the validity of the overlapping conditions, it would lend them a degree of empirical support that they currently lack, and thereby strengthen their claim to scientific status. Without experimental and clinical controls, however, it is difficult to rule out equally credible alternative explanations of therapeutic improvement (Erwin, 1985, 1993).

---

2This demand is congruent with what Erwin (1997, p. 75) calls the principle of differentialness in science: “Data confirm a hypothesis only if they do so differentially. . . . For any body of data D and hypothesis H, D confirms H only if D provides some reason for believing that H is true, and does not provide equal (or better) reason for believing some incompatible rival that is just as plausible.”
Here, I will explore only one alternative explanation of the phenomena that occur in the insight-oriented psychotherapies: the therapeuticty of insight-mimicking explanatory fictions and therapy-induced self-deceptions and illusions. In some respects, this explanation is at once the most obvious and the most easily overlooked. This is because it is so clearly at odds with the standard view. Rather than beginning with the obvious assumption that the insight-oriented psychotherapies are likely to generate veridical insight, and that any failures to do so are to be explained as failures in the application of the treatment methods (rather than failures in the methods themselves and the theories that explain them), I will begin with the opposite assumption: that the insight-oriented psychotherapies are likely to generate illusions, deceptions, pseudo-insights, and adaptive self-misunderstandings that convincingly mimic bona fide insight; and that their failures to generate bona fide insight are not merely failures in the application of the treatment methods but failures in the methods and the therapeutic theories themselves.

This oppositional strategy is based on two guiding assumptions, one drawn from philosophical psychology and one drawn from social and political philosophy. The first assumption is that there are many more ways to be wrong, deceived, or deluded about oneself—one's psychology, motivational structure, feelings, life history, and character traits—than there are ways to be knowledgeable. Given the self’s complexity and open-endedness, it is easy to be mistaken about such things as: the psychological centrality, developmental history, and causal role of the parts of the self vis-à-vis one another; the role of character traits in the formation of action; and the relative mismatch or coherence of the different parts of the self.

The second guiding assumption is that it may be more socially and psychologically adaptive to remain deceived, ignorant, or deluded about oneself than to be knowledgeable; that is, to live with an understanding of self that is shaped by misinformation, fabrication, false theories, and benign fictions. The idea that humans are subject to illusions because they are subject to illusions requires illusions is a theme found throughout the writings of Spinoza, Marx, Nietzsche, Heidegger, and others. Pseudo-interpretations of human experience, behavior, and the self nonetheless serve the function of interpreting suffering, unexplained natural forces, and puzzling behaviors, thereby making them more bearable than they would be otherwise, and supplying for them socially sanctioned remedies. This is most clearly seen in shamanistic and religio-magic healing rites, which interpret human powerlessness against a mysterious and threatening natural world and which supply imaginary remedies that give a semblance of power over alien forces. But similar pseudo-interpretations serving many of the same purposes, and pressing into service many of the same cognitive and imaginative competences, are at work in contemporary psychotherapeutic settings. The cultivation of the illusion of gaining special insight into the forces that govern the human mind and human behavior, and of having risen above the condition of powerlessness, is psychologically palliative even if it leaves the real state of powerlessness intact. It diverts attention from the facts of suffering and powerlessness, and defuses social discontent.

I use the generic term placebo insights to describe the range of therapeutically induced pseudo-insights, self-deceptions, explanatory fictions, and adaptive self-misunderstandings that clients in the insight-oriented psychotherapies mistakenly interpret as valid forms of self-knowledge and insight. Of course, not all therapeutically induced pseudo-insights and illusions are placebogenic. It is important to distinguish between those that have positive effects, those that have negative effects (or are “nocebogenic;” Hahn, 1997), and those that have no discernible therapeutic effects at all. Moreover, there are many other alternative explanations of the phenomena that occur in the insight-oriented psychother-
Placebo Insights

Placebo insights are explanatory fictions and therapy-induced self-deceptions that convincingly mimic bona fide insights. They are constituted by observations and explanations that appear to be authentic, but in fact they have no more explanatory power and descriptive validity than psychological sugar pills. Despite their explanatory poverty, placebo insights remain convincing to clients, whose capacities for independent critical judgment are under constant pressure from strong expectancy effects regarding insight acquisition, and from high motivational levels to reduce the cognitive dissonance occasioned by the therapeutic encounter. Placebo insights carry a degree of coherence and plausibility that is marginally greater than the pretherapeutic interpretation of the target disorders that serves as the baseline against which they are compared. They appear to effect therapeutically beneficial symptom relief or intrapsychic personality change, and thus appear to display genuine causal agency. But the effectiveness of placebo insights is unrelated to their content and truth-value. They are effective only to the extent that (a) they are subjectively pleasing to the clients who have struggled to acquire them,\(^4\) and who expect their struggles to pay off; (b) they receive consensual support, and are consistent with the existing knowledge structure that is supplied by the therapeutic theory; and (c) they appear to refer to the salient aspects of the client’s psychological make-up and behavior, as these are identified under the therapeutic theory’s postulated hypothetical constructs. The therapeutic effectiveness that placebo insights appear to manifest can be explained adequately in terms of factors that have little to do with truth value.

Placebo insights are a direct function of the highly volatile nature of the therapeutic encounter in exploratory psychotherapy. Given the psychological and cognitive pressures generated by the therapeutic encounter (Calestro, 1972; Strupp, 1972), including the wide range of expectancy effects, and the occurrence of doctrinal compliance (Grünbaum, 1986, 1993), there is a permanent risk that clients may be persuaded to accept from their therapists as true certain putatively truth-tracking interpretations that are in fact psychologically and historically incorrect (Mendel, 1964; Schmideberg, 1939). Epistemically, this leads to a downward spiral. Under the influence of false or inexact interpretations, leading questions (Fish, 1986), and subtle cues, clients will make false discoveries, and will come to formulate for themselves false insights on the basis of evidentiary and interpretive criteria that have been progressively weakened by continued exposure to the pressures of the therapeutic situation and the therapist’s theoretical orientation.

It would be a mistake to assume that placebo insights are easily identifiable from among the range of insights yielded by the exploratory psychotherapies. It also would be a mistake to assume that placebo insights would be winnowed out by a kind of rational self-corrective process that is intrinsic to a properly executed and methodologically scrupulous therapy. Just as therapeutic suggestion is considerably more subtle than crude persuasion that is directed against symptoms (e.g., “your anxiety will abate if you count to 50”), so placebo insights are considerably more subtle than simple-minded falsehoods or fanciful fictions. They are to be distinguished, therefore, from more obvious forms of psychological jargon (or “psychobabble”), those forms of therapeutic discourse that

---

\(^3\)One exception is Grünbaum (1986, 1993), who discusses the role played by pseudo-insight in classical Freudian psychoanalysis. Grünbaum does not, however, give an analysis of the concept of pseudo-insight.

\(^4\)Shapiro (1971) points out that the term placebo is derived from the Latin verb “placere,” meaning “to please.”
characteristically smooth over psychological and behavioral complexity with vague and portentous terms that are stretched to the point of vacuity (Rosen, 1975). Placebo insights, unlike insights constituted by simple-minded psychological jargon, are characterized by a certain level of psychological robustness and complexity; and they are influenced by the therapist’s technical terminology, which has the power to seduce clients who are psychologically minded and mystify those who are not. Thus, Shapiro (1971, p. 459) characterizes placebogenic psychotherapies as “elaborate, detailed, expensive, time-consuming, fashionable, esoteric, and sometimes dangerous.” The identifiability of placebo insights is made more difficult by the fact that prolonged exposure to a treatment method and theoretical orientation tends to sensitize clients to the importance of certain selected issues and explanatory strategies, which then predisposes them to expect that these issues will be addressed only by certain methods of exploration and certain kinds of insights.

It also would be a mistake to think that the acquisition of placebo insights leaves clients unchanged in all but their beliefs about themselves and their disorders. Placebo insights are not without psychological and behavioral consequences. They are not false in the way that false beliefs about states of affairs in the world are false. False beliefs about the chemical composition of a brick of gold, for example, do not alter the gold itself or the kinds of evidence it yields. By contrast, importantly false beliefs about one’s own psychology, behavior, or life history have the potential to alter one’s actual self. With sufficient reinforcement, importantly false beliefs can become just as powerful a determinant of behavior, attitude, and character as veridical beliefs.

There are a number of ways in which the changes occasioned by the acquisition of the false beliefs constituting placebo insights take place. One of the more philosophically interesting instances of change occurs when the treatment methods to which clients are subjected in the insight-oriented psychotherapies generate some of the very psychological and behavioral facts that the clients claim to “discover” in their explorations, and which the clients interpret as having existed prior to the intervention of the therapy: when, for example, the treatment methods of classical Freudian psychoanalysis generate psychological, behavioral, and phenomenological clinical material that appears to indicate the presence in clients of Oedipal complexes or castration anxieties; when the treatment methods of Jungian analysis generate clinical material that suggests the presence in clients of the workings of the collective unconscious; or when the treatment methods of Adlerian therapy generate clinical material that suggests the presence in clients of masculine strivings and feelings of inferiority (Marmor, 1962, 1970). These are cases of “therapeutic conformability;” that is, the surreptitious creation in clients, through the agency of powerful therapeutic treatment methods and the reflexive feedback loops they generate, of an overlay of psychological, behavioral, and notional artefacts that mimic the existence of the entities to which the therapeutic theory refers. Therapeutic conforming occurs when clients grow into and conform to the therapeutic modalities that have grown around and conformed to them. Psychotherapies that display this peculiar characteristic are to a certain extent self-confirming, and thereby less than scientific (Farrell, 1981).

A simple example of therapeutic conformability occurs in Freudian dream analysis. Clients who have been sufficiently exposed to the treatment methods and therapeutic expectations of dream analysis begin to produce dreams—or memories of dreams—that fit the models and expectations of the very dream analysis they are undergoing. This is not an uncommon phenomenon. The dreams, which appear to clients to be spontaneous and naturally produced, are in actuality therapy-induced artefacts that would not have occurred without exposure to the treatment methods of the therapy. What appears to be
uncontaminated clinical material is then subjected to the same treatment methods and interpretive strategies that first generated it, which in turn generates more artefactual dreams, and so on. The feedback loop this creates comes to constrain and guide the kinds of insights that clients develop about the meanings of their dreams, and the kinds of insight-driven changes they experience. Their insights, which have the potential to influence behavior, attitude, and character, are in actuality therapeutic artefacts generated by the treatment methods of the therapy.

This is a fairly simple example. If therapeutic conforming were restricted only to simple and easily identifiable cases such as this, then it would not constitute a serious threat to the scientific status of the insight-oriented psychotherapies nor, a fortiori, to the core principles of the standard view. Such cases could be regarded merely as side effects of the treatment methods—harmless by-products that could, at least in principle, be adequately controlled for. But the phenomenon is not restricted in this way, and its presence in other forms of psychotherapy is more insidious. Consider how it plays out in existential psychotherapy, where clients who have been sufficiently exposed to the therapy and the therapeutic theory begin to manifest behaviors, thoughts, and attitudes that display a distinctive existential bearing. In some cases, clients who never before had experienced “ontological moods” begin to report feelings of existential anxiety or despair, thereby confirming the expectations of the existential therapists that their clients’ pretherapeutic problems were incipiently existential in nature. In such cases, the therapy confers a unified morphology upon the otherwise diffuse, inchoate, or anomalous pretherapeutic problems that first prompted the clients to seek help. What to clients and therapists appears to be authentic discovery is actually an artefact of the therapeutic situation that would not have been encountered independently of the therapy. But the therapeutic interference does not stop here. When clients begin to display behaviors and to have experiences that conform to the expectations of the treatment methods, existential therapists are naturally encouraged to proceed with their therapeutic strategies on the grounds that the clinical material confirms the validity of the existential approach. As in dream analysis, this creates a feedback loop, with clients giving back to therapists responses and behaviors that increasingly bear the imprint of the therapeutic modality (Farrell, 1981). Naturally, some therapeutic artefacts go against the “grain” of the clients’ pretherapeutic psychological, behavioral, and phenomenological identities, thereby creating an unstable notional overlay that is poorly fitted to the facts.

The presence of therapeutic conformability in the insight-oriented psychotherapies suggests that the four central principles of the standard view of insight-oriented psychotherapy tend to oversimplify and overidealize matters. Much more is going on in the insight-oriented psychotherapies than the straightforward exploration and discovery of antecedently given facts: There also is a dynamic interchange between the treatment method and the “material” upon which it operates—a dialectic of forming, conforming, and re-forming. It is not as if the facts about the client’s soul are there awaiting exploration, and will be discovered only when accessed by a precisely applied method that fits them snugly like a key opening a lock; rather, the lock is openable because it is made to conform to the specific morphology of the key that is applied to it.

If clients had been subjected to other therapists working with different methods and theoretical frameworks, then the therapy-induced artefacts the clients would “discover” in their explorations would have been different; but there would have been no less of an artefactual overlay of some sort. It is not as if different insight-oriented psychotherapies are engaged in exploring and converging upon the same determinate facts from slightly different angles. Rather, different therapeutic approaches generate different artefactual overlays, each of which alters clients in ways that cannot subsequently be factored out.
Most of this therapeutic formation occurs surreptitiously, without the explicit awareness of the clients to whom it is happening. While clients interpret the changes they experience as a matter of coming to a clearer awareness of certain salient and antecedently given facts about their psychology, behavior, and life history, what is really happening is that the therapeutically driven changes in their lives are making their forms of self-awareness clearer and more determinate than they would otherwise have been, by generating the very objects of that self-awareness.

The metaphor of insight-oriented psychotherapy as a kind of mirror of the soul, or as a kind of archaeology, is therefore misleading. Clients are less like archaeologists of the soul and more like explorers in a strange land whose every step forward alters the landscape which they are exploring.

Because of the spurious relation between theory and empirical fact that it occasions, therapeutic conformability is one among a range of phenomena which impugns the scientific status of the insight-oriented psychotherapies. Generally, the scientificity of a theory is a function of its testability by empirical observations that are derived from the theory. Empirical testing results in the confirmation or disconfirmation of the theory, and it serves to ensure the ongoing rational self-correction which is the hallmark of scientific method. Some theories (such as astrology) resist testing altogether; others resist testing of the right sort. Theories comprising the latter group are not without empirical content, but the content they yield is not produced using suitably independent measures. The evidence, observations, and predictions that in testing are brought to bear on such theories are themselves a function of the theories they putatively test in the sense that they would not exist independently of the application of the theories. Such theories—including those theories in the insight-oriented psychotherapies that are vulnerable to therapeutic conformability—are self-confirming (rather than independently confirmable or disconfirmable). One of the drawbacks with self-confirming theories in the psychotherapies is empirical uncertainty. Without both clinical and experimental controls, psychotherapists never can be sure to what extent the clinical material to which they turn to validate their theories is the result of the interference of the very treatment methods that elicit the clinical material in the first place, or material which could be acquired independently.

Therapeutic Placebos

Why has the placebogenic character of psychotherapeutic insight and interpretation—and more broadly, the therapeuticity of explanatory fiction—been overlooked so widely in the theory and practice of the insight-oriented psychotherapies? One reason is that placebos and placebo effects in all healing modalities are generally regarded in poor light. To many they imply that suggestion, deception, or coercion are being pressed into service by manipulative therapists without regard to the autonomy of clients. The deliberate use of placebo therapy has been criticized as irresponsible medical practice, as tantamount to charlatanism, and as unscientific.

From a historical point of view, however, these criticisms are in the minority. The principle of respect for patient autonomy, and the emphasis on transparency and truth telling, are historically recent ideas, with origins in the Kantian commitment to the universal duty of veracity (Crenshaw-Rawlinson, 1985). The deliberate use of placebos in medical treatment and the practice of benevolent deception, by contrast, have ancient roots. Both Hippocrates and Galen recognized the healing power of symbolic elements in patient–doctor relations, and acknowledged the importance of protecting patients from knowledge (and self-knowledge) that could be harmful (Lain Etralgo, 1970). Hippocrates (1979), for instance, argued that physicians must practice their art
calmly and adroitly, concealing most things from the patient while... attending him. Give encouragement to the patient to allow himself to be treated, turning his attention away from what is being done to him; sometimes reprove sharply and emphatically, and sometimes comfort with solicitude and attention, revealing nothing of the patient’s future or present condition. (p. 297–299)

Socrates also recognized the healing power of rhetoric, fiction, and persuasion. Observing the soul doctors of the day, he claimed that “the cure of the soul has to be effected by the use of certain charms, and these charms are fair words” (Plato, 1961; also see Gill, 1985).

More recently, Pierre Janet used therapeutic deception, explanatory fictions, and the psychological equivalent of sugar pills to great effect, curing his patients by fabricating stories about their condition and persuading them to believe that they were true. His therapeutic methods included hypnosis, suggestion, “monoideism,” moral education, guided imagery, and the deliberate manipulation and reconstruction of the patient’s memories (Ellenberger, 1970). Janet’s directive techniques stand in sharp contrast with Freud’s ostensibly nondirective exploratory techniques. Where Freud was convinced that it was the patient’s talking that cured, Janet was convinced that it was the patient being talked to that cured (Borch-Jacobsen, 1996); and where Freud was convinced that he had found a method that cured by enabling patients to discover the truth about their pasts and their unconscious motivations, Janet knew that he was deliberately misleading his patients in order to cure them (Hacking, 1995; Janet, 1925, p. 339).

Of course, the claim that the insight-oriented psychotherapies produce interpretations and insights that are the equivalent of psychological sugar pills raises a broader question: What precisely is a placebo, and how is it to be distinguished from a nonplacebo?

Shapiro and Morris (1978, pp. 371–372) characterize a placebo as any therapy or component of therapy that is deliberately used for its nonspecific, psychological, or psychophysiological effects, or that is used for its presumed specific effect, but is without specific activity for the condition being treated. . . . Specific activity is the therapeutic influence attributable solely to the contents or processes of the therapies rendered. The criterion for specific activity (and therefore the placebo effect) should be based on scientifically controlled studies.

This is a good first approximation, but it leaves a number of questions unanswered. For example, what is specific activity? How is specific activity to be distinguished from nonspecific activity? And what is the role of the therapist’s intentions?

Grünbaum’s (1994) analysis of the concept of placebo helps to sharpen the definition given by Shapiro and Morris. First, Grünbaum notes that because the effects of placebos (and placebo insights) can be defined just as sharply as the effects of nonplacebos, the term incidental to characterize the relevant placebo components of a therapy should replace the term nonspecific, which connotes indistinctness. Second, he argues that there is a distinction between intentional and inadvertent placebos because both patients and physicians can be unaware that the treatments being administered are placebos. Something counts as an intentional placebo if (a) none of the characteristic constituents of the therapy (as specified by the therapeutic theory) are remedial for the disorder, (b) if the therapist believes that none of the specific constituents are remedial, and (c) if the therapist believes that the treatment method is remedial by virtue of other incidental aspects of the treatment. By contrast, something counts as an inadvertent placebo if (a) none of the characteristic constituents of the therapy are remedial for the disorder, (b) if the therapist believes that the specific treatment method is remedial for the disorder by virtue
of its characteristic constituents (as specified by the therapeutic theory), and (c) if the patient believes that the treatment method is remedial for the disorder by virtue of its characteristic constituents.

Grünbaum’s analysis of Shapiro and Morris’s definition of placebo is useful for understanding the dynamics of placebo insight. An insight-oriented psychotherapy is characterized by a therapeutic theory (e.g., the theory of Freudian psychoanalysis) that recommends a treatment method for a particular target disorder insofar as that disorder is identified under the nosological component of the theory (e.g., neurotic obsession). The treatment method contains both characteristic factors (e.g., analysis, free association, analytic interpretations) and incidental factors (e.g., the payment of a high fee for the therapy, which may be a catalyst for the patient’s receptivity to treatment). Veridical insights count as one of the characteristic factors of the therapy; placebo insights, by contrast, count as one of the incidental factors. An insight-oriented psychotherapy counts as a nonplacebo if one or more of the characteristic factors have a positive remedial effect on the disorder. It counts as a placebo therapy if none of the characteristic factors are remedial for the disorder, but the disorder is nonetheless improved because of the effect of the incidental factors. In classical Freudian psychoanalysis,

the patient’s correct, affect-discharging insight into the aetiology of his or her affliction is the one quintessential ingredient that distinguishes the remedial dynamics of his treatment modality from any kind of treatment by suggestion. Treatments by suggestion, [Freud] charged, leave the pathogenic repressions intact, and yield only an ephemeral cosmetic prohibition of the symptoms (Grünbaum, 1994, p. 295).

Janet clearly relied upon deliberate placebos to effect therapeutic change. He was well aware of the role of fabrication and deception in psychological healing, and the central importance of patients believing that the diagnoses and interpretations he supplied them with, and the insights they acquired as a result of the treatment, were true. The standard version of the insight-oriented psychotherapies, by contrast, does not adequately acknowledge the central role played by the presence of explanatory fictions, and therapy-induced self-deceptions and self-misunderstandings. It is therefore vulnerable to inadvertent placebogenesis precisely because of its commitment to the four core principles and to the principle of respect for client autonomy. To the extent that the standard version holds that the treatment methods of the insight-oriented psychotherapies are remedial for target disorders as specified by the relevant therapeutic theories, it fails to acknowledge the presence of therapeutically induced fictions and illusions informing the client’s insights (see Hacking, 1995, pp. 195–196). Some exploratory psychotherapies in the postmodern tradition try to avoid this problem altogether by claiming that there is no such thing as truth, but this is tantamount to abandoning all hope of progress in understanding the nature of human well-being, suffering, and health (Held, 1995).

It is clearly a gross exaggeration to claim that placebo insights are the only kind of insights generated in the insight-oriented psychotherapies. Clients sometimes track the truth in their explorations, and sometimes hit upon the truth. In some cases, this is an accidental occurrence that has no significant causal connection to the presumed activity of the therapy; in other cases, it is a direct function of the combined effects of the characteristic and incidental factors of the therapy. Moreover, some insights may have elements of both truth and falsity. Because of the absence of sufficiently rigorous experimental and clinical controls, however, it is often difficult in actual practice to distinguish the incidental and placebogenic factors from the characteristic factors of the insight-oriented psychotherapies. The interspersion of the incidental factors with the characteristic factors of a therapy is perhaps best represented as a unitless sliding scale that measures overall
therapeutic improvement. At one end of the scale, therapeutic improvement is directly attributable to factors that are characteristic of the treatment method: The psychoanalytic method of free association, for example, operates directly upon the specified target disorder like a key opening a lock and occasions veridical insights that trigger substantive therapeutic improvements. Nearing the other end of the scale, therapeutic improvement is attributable primarily to the agency of incidental factors: A number of different keys (and a number of alternative insights) would have done equally well in opening the lock (or opening it by making the lock conform to the contours of the key). At the very far end of the scale, the characteristic features of the treatment method, including the content of the therapist’s interpretations and the client’s insights, are epiphenomenal. Whatever insights clients acquire in psychotherapy are a kind of cognitive window dressing, even if, subjectively, clients are convinced of their depth and value.

Moreover, in actual practice it is often difficult to distinguish bona fide insights from placebo insights. Even if it is the case that there are many more ways to be wrong, deceived, or deluded in the exploration of the soul than there are ways to be knowledgeable, and that the tendency to err is greater in the insight-oriented psychotherapies than in nontherapeutic practices, it does not follow that there are no realizable epistemic conditions in which veridical insight is achievable. The chances of establishing the appropriate conditions under which veridical insights can be acquired are improved (a) as clients gradually reduce their dependence on treatment methods that are at risk of suggestion, placebo and expectancy effects, therapeutic conformability, and other forms of epistemic contamination; (b) as clients increase their awareness of the prevalence of explanatory fictions and therapy-induced self-deceptions in therapy; and (c) as clients press into service appropriately targeted prophylactic and decontamination measures (e.g., knowledge of common logical fallacies and errors in human reasoning), with a view to systematically reducing the chances of falling into the more significant forms of error (see Jopling, 2000, for a discussion of the epistemic standards for self-knowledge).

From the client’s point of view, however, this fallibilist stance to insight-oriented psychotherapy is difficult to realize. First, it commits clients to living with the uncertainty that their attempts to sift out all forms of epistemic contamination can never be finalized. In the end, clients in the talking cures can never know with certainty to what extent their newly acquired insights are therapeutic artefacts or bona fide acquisitions. Second, the fallibilist stance does not coincide with the primary goals of the talking cures. Clients are not epistemologists, and epistemology is not one of their primary concerns when they commence what they consider to be a journey of exploration of the soul. Warnings of the threat of epistemic contamination can be expected to leave clients unmoved, especially if the removal of the threat does not contribute immediately to what they consider to be therapeutic progress. Moreover, the introduction into the therapy of a set of second-order epistemic concerns can be countertherapeutic, insofar as it deflects their attention away from the issues that first prompted them to enter therapy, and interferes with the therapeutic exclusivity that is one of the crucial factors at work in all healing modalities. A certain amount of selective inattention to second-order issues is essential while the psychotherapy is in progress, just as selective inattention to alternative artistic styles is essential while artists are in the process of creating an artwork. A narrow focus is therapeutically beneficial, as is a commitment to the therapy that is not reflective about its own epistemic conditions of possibility.

Those who defend the standard view can counter the placebo insight hypothesis with two arguments. First, any exploratory psychotherapy that generated therapeutic change through placebo insights would prove to be ineffective because the insights would be poorly fitted to the contours of the client’s life processes and inadequately action-
guiding. Second, placebo insights can provide only temporary symptom relief, and clients will relapse after the initial effects of insight wear off.

Both counterarguments underestimate the power of placebo effects in the insight-oriented psychotherapies, and, more broadly, both underestimate the role of imagination, fiction, and illusion in psychological healing. Placebo insights are effective and durable to the extent that they are adaptive and oriented to problem solving. The high level of functionality of explanatory fictions receives support from a number of quarters in philosophy and psychology, where illusion, deception, blind stupidity, oversimplification, and falsification are recognized as life-enhancing virtues, and as conducive to effective agency, and where the unqualified commitment to truth has been regarded as destructive of the blind forward momentum of life.

Nietzsche, for instance, argued that human beings could not survive without limited perceptual, affective, and cognitive horizons: a “narrowing of our perspective, and thus in a certain sense stupidity, [is] a condition of life and growth” (Nietzsche, 1886/1966). The plays of Henrik Ibsen (1884/1961) and Eugene O’Neill (1946) also illustrate how existence for the average person is tolerable only with a veil of comforting and self-serving illusions that filter out the harsher elements of life. Without the psychological crutches supplied by robust pipe dreams and “life-lies,” it would be difficult to function normally and to maintain adaptive self-regarding attitudes such as self-esteem and self-contentment. The plays of Ibsen and O’Neill thus focus upon a recurrent theme: When pipe dreams and life-lies are burst, in the name of an honest confrontation with reality, tragedy and despair inevitably follow.

Recent work in cognitive and social psychology bears out these philosophical and literary insights. One area of research defends the view that persons generally tend to be mistaken in their self-reports and self-descriptions, and generally tend to rationalize or confabulate according to empirically underdetermined socially embedded schemas when asked about their motives, desires, and character traits (Nisbett & Wilson, 1977). Another area of research defends the view that positively biased “creative” illusions about the self play a more significant role in the maintenance of mental health, as well as in the maintenance of caring interpersonal relations and a sense of well-being, than do accurate self-perceptions and veridical self-knowledge. These illusions include unrealistically positive self-evaluations, exaggerated perceptions of personal control, and unrealistic optimism about the future (Jopling, 1996; Taylor, 1989; Taylor & Brown, 1988).

If these claims are valid, then veridical self-knowledge is neither an indispensable ingredient of well-being nor an essential component of functioning successfully in complex social environments. By pressing into service the client’s receptivity to the therapeutic agency of explanatory fictions, the insight-oriented psychotherapies are cultivating adaptive illusions under the guise of helping clients gain veridical insight. The therapeutically of explanatory fictions and therapy-induced self-deceptions is a hard-to-avoid consequence of the ideal of pursuing the truth.

It would be mistaken to conclude, therefore, that placebo insights are always and necessarily bad for the insight-oriented psychotherapies, and ought to be eradicated at all costs if psychotherapy is not to be reduced to the level of charlatanism or pseudoscience. Placebo insights help clients feel that they are more insightful, more coherent, and more in touch with themselves than they would otherwise have had occasion to be, even if there is a clear sense in which they are significantly less in touch with themselves than they believe themselves to be. Moreover, placebo insights have the potential to serve the purely instrumentalist function of offering up useful guides or tools for post-therapeutic self-inquiry, which might ultimately lead to genuinely truth-tracking insights—or to further insight-mimicking illusions and deceptions. But if placebo insights can be said to
“work” for clients, they do not work for the reasons that clients and therapists think they work, that is, because they get at the truth, and because they supply accurate explanations of the client’s target disorders and psychological make-up. Rather, they work to the extent that they bring about a more fictionalized contact with reality, including the reality of the self. Because how things seem to the client and the therapist and how things are are quite distinct, the scientific explanation of how the insight-oriented psychotherapies work is significantly different from that provided by therapists in their theoretical accounts of their practices and from that provided by clients in their first-person reports about their newly won insights and newly changed behaviors.

Just as one would be mistaken to conclude that placebo insights are bad for the theory and practice of insight-oriented psychotherapy, one also would be mistaken to conclude that they ought to be fully embraced as legitimate tools, as Janet and others have argued.

First, if placebo insights are admitted into therapeutic practice as useful, then it becomes difficult to see how the legitimate uses of insight-oriented psychotherapy can be demarcated from the illegitimate uses. The criterion for demarcating insight-oriented psychotherapies that trade in fiction and fantasy from those that track the truth—and, a fortiori, for demarcating pseudoscientific from scientific psychotherapies—is much less clear and much less comprehensive that what defenders of the standard model would like.

Second, the benefits of placebo insights always must be weighed against the undesirable interpersonal consequences that may follow from their cultivation (Jopling, 1996). If one of the central functions of placebo insights is to mediate the harsh impact of reality through explanatory fictions, then one of the potential consequences is the editing or filtering out of negative social feedback—that is, the criticisms, advice, disapprovals, and hurt feelings of other persons. But while this may be beneficial from a strictly self-regarding standpoint, it can result in a kind of “other-blindness:” that is, a systematic misrecognition or misinterpretation of the effects of one’s actions upon others. If therapeutically induced illusions and deceptions leave clients less responsive to negative social feedback, then they are less capable of experiencing the kinds of moral and personal growth that are only possible in contexts of interpersonal relations unencumbered by illusions and deceptions.

Finally, embracing explanatory fictions and therapy-induced self-deceptions for their therapeutic value runs into conflict with some of our deeper convictions about what it is to be a fully developed and fully self-aware human being. Generally, knowledge is preferable to ignorance. Self-knowledge is widely regarded as one of the goods of human life while self-ignorance and self-deception are regarded as moral failings. Part of our conception of what it is to be a responsible moral agent is that we place a high value on being cognizant of who we are, where we are going in our lives, what moves us to action, how we affect others, and how others regard us. This amounts to more than a superficial awareness of self. We ascribe genuine self-knowledge to those who have raised the question “Who am I?” in a searching and fundamental manner; who have assumed a stance of reflective self-evaluation and self-criticism toward core desires, beliefs, motives, emotions, and character traits; and who have achieved a significantly greater degree of action-guiding insight with respect to their motives and values than they would otherwise have had occasion to achieve (Jopling, 2000). With this increased awareness comes a greater degree of responsibility. Any attempt to reduce this level of awareness of self, and the self’s relation to others, through the mediation of explanatory fictions and therapy-induced illusions is like an act of disconnection that severs the self from itself. Embracing placebo insights uncritically for their therapeutic value carries with it the implication that it is better to see the world (and oneself) through a veil of illusion or fantasy for the deluded contentment it delivers, than it is to see the world as it is.
References


